

## COMPARAÇÃO DE DUAS METODOLOGIAS DE AMOSTRAGEM ATMOSFÉRICA COM FERRAMENTA ESTATÍSTICA NÃO PARAMÉTRICA

**Maria João Nunes\*** e **Maria Filomena Camões**

Departamento de Química e Bioquímica, Faculdade de Ciências, Universidade de Lisboa e Centro de Electroquímica e Cinética da Universidade de Lisboa, Edifício C8, Campo Grande, 1749-016 Lisboa - Portugal

**Francis McGovern**

Environmental Protection Agency, St. Martins House, Waterloo Road, Dublin 4, Ireland

**Sebastião Santos e Frank Raes**

European Commission, Joint Research Centre, TP 290, Ispra (VA), I-21021, Italy

Recebido em 3/6/03; aceito em 22/9/04; publicado na web em 28/1/05

COMPARISON OF TWO ATMOSPHERIC SAMPLING METHODOLOGIES WITH NON-PARAMETRIC STATISTICAL TOOLS. In atmospheric aerosol sampling, it is inevitable that the air that carries particles is in motion, as a result of both externally driven wind and the sucking action of the sampler itself. High or low air flow sampling speeds may lead to significant particle size bias. The objective of this work is the validation of measurements enabling the comparison of species concentration from both air flow sampling techniques. The presence of several outliers and increase of residuals with concentration becomes obvious, requiring non-parametric methods, recommended for the handling of data which may not be normally distributed. This way, conversion factors are obtained for each of the various species under study using Kendall regression.

Keywords: atmospheric sampling; aerosol analysis; non-parametric statistics.

### INTRODUÇÃO

Amostragens atmosféricas têm como inerente o facto de o ar se encontrar em movimento devido à circulação das massas de ar. No caso particular dos aerossóis atmosféricos, o fluxo de volume de ar amostrado é particularmente determinante, pois pode conduzir à recolha de amostras com composições diferentes, devido aos vários tamanhos de partículas e à diferente acção de sucção do ar amostrado.

No âmbito do Projecto "Long-term Aerosol Characterization Experiment-2" foi desenvolvido um programa de monitorização de aerossóis na região do Atlântico Norte de novembro de 1995 a Janeiro de 1998, entre 44°N a 23°N e 25°W a 8°W, complementado com uma campanha intensiva em junho - julho de 1997. Durante o Projecto, a amostragem de aerossóis com dimensões inferiores a 10 µm foi efectuada com amostradores de elevado fluxo de volume de ar. Na campanha intensiva foi simultaneamente utilizada uma metodologia de baixo volume de ar que se pretendeu integrar e validar permitindo a comparação dos resultados de ambas. O objectivo deste estudo torna-se particularmente relevante quando se utilizam dados de referência obtidos com metodologias de amostragem diferentes, nomeadamente no caso da modelação atmosférica em que se utilizam dados obtidos em programas de monitorização de concentrações atmosféricas de aerossóis. Simultaneamente também se verifica que no caso de algumas espécies iónicas os resultados obtidos com diferentes metodologias de amostragem apresentam uma grande dispersão de valores com desvios da normalidade, o que impede a sua comparação e a obtenção de factores de conversão utilizando estatística paramétrica.

As estações de monitorização estavam equipadas com sensores meteorológicos e com contadores de núcleos de condensação (número de partículas, CN). A amostragem foi controlada por computa-

dor com condições pré-definidas da variação de CN, direcção e velocidade do vento. Na estação de Porto Santo, Ilha da Madeira, Portugal, localizada no Nordeste do Oceano Atlântico, procedeu-se à amostragem de aerossóis utilizando amostradores de alto e de baixo fluxo de volume de ar.

Considerando os compostos químicos de interesse, os tipos de amostradores disponíveis e o tipo de filtro utilizado, é necessário encontrar as respostas para as seguintes questões:

- Quais os procedimentos que deverão ser controlados para uma comparação bem sucedida?
- As diferenças observadas nos resultados são devidas à amostragem ou às variantes de procedimento na determinação analítica?
- Como é que podem ser identificadas pequenas diferenças nos resultados devido a diferentes metodologias de amostragem?

A concentração atmosférica da fracção de aerossol solúvel em água foi determinada por cromatografia iónica (CI) para as espécies iónicas maioritárias. A utilização de ferramentas não paramétricas para comparação das duas metodologias de amostragem foi efectuada para todas as espécies iónicas objecto deste trabalho. Foram obtidos factores de conversão dos resultados de metodologia de amostragem de elevado volume em baixo volume.

### MÉTODOS ANALÍTICOS

A amostragem foi controlada por computador através de dados obtidos do anemómetro e do contador CN. Cada amostrador estava ligado a uma entrada PM10 (Matéria Particulada 10 µm), tendo sido recolhidos aerossóis de dimensões inferiores a 10 µm. A amostragem de elevado volume de ar foi realizada com filtros Whatman 41® (φ = 150 mm), a um fluxo de ar de 0,5 m<sup>3</sup> min<sup>-1</sup>. A amostragem de baixo fluxo de ar foi realizada num suporte de dois andares, com filtros Nuclepore® e Whatman 41® (φ = 47 mm), a um fluxo de 0,016 m<sup>3</sup> min<sup>-1</sup>.

\*e-mail: mariajnunes@hotmail.com

Foi efectuada a análise quantitativa da fracção de aerossol solúvel em água. Os procedimentos analíticos de extracção em água foram diferentes devido ao facto de as concentrações nos filtros provenientes das duas metodologias de amostragem serem substancialmente diferentes. A quantidade de amostra analisada e os métodos de extracção com água desionizada (18,2 MΩ cm) foram diferentes; nas amostras de alto volume foi utilizado 1/4 de filtro e a extracção foi efectuada por centrifugação e nas amostras de baixo volume foram utilizados os dois filtros e a extracção foi efectuada em banho ultra-som.

Os extractos foram analisados por CI para as espécies maioritárias catiónicas (Na<sup>+</sup>, NH<sub>4</sub><sup>+</sup>, K<sup>+</sup>, Mg<sup>2+</sup> e Ca<sup>2+</sup>), aniónicas (Cl<sup>-</sup>, NO<sub>3</sub><sup>-</sup> e SO<sub>4</sub><sup>2-</sup>) e anião de ácido orgânico (MSA, CH<sub>3</sub>SO<sub>3</sub><sup>-</sup>). Foi então determinada a massa iónica em cada filtro (µg). A concentração atmosférica foi obtida considerando o volume de ar amostrado.

## TESTES ESTATÍSTICOS NÃO-PARAMÉTRICOS

A estratégia adoptada para a análise de resultados foi a utilização de métodos estatísticos não-paramétricos, inerente ao facto de algumas espécies apresentarem grande dispersão de valores. Os modelos não-paramétricos implicam ordenação ascendente das variáveis, sendo suas principais vantagens as variáveis não dependerem da distribuição da população e não serem afetadas pela presença de valores aberrantes. Na análise estatística foram utilizados o teste de ordenação de valor absoluto de Wilcoxon, o coeficiente de correlação de Spearman e a regressão de Kendall<sup>1-4</sup>.

O teste de ordenação do valor absoluto de Wilcoxon compara dois grupos emparelhados. A diferença entre cada par de observações é calculada analisando a série de diferenças entre  $n$  pares de observações, assumindo que as diferenças estão distribuídas simetricamente em relação ao seu valor médio. Estas diferenças são ordenadas por ordem crescente do seu valor absoluto e o sinal positivo ou negativo é atribuído ao valor absoluto. Em seguida são efectuadas as somas da série ordenada positiva e da série ordenada negativa. O somatório com menor valor absoluto é denominado  $T_s$ . Para um número de amostras,  $n$ , maior que 50, tal como o utilizado neste trabalho, o valor estatístico  $t_s$  é calculado através da Equação<sup>4</sup>:

$$t_s = \frac{T_s - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1/2)(n+1)}{12}}} \quad (1)$$

$t_s$  é comparado com o valor crítico,  $t_c$ , utilizando uma tabela estatística de distribuição- $t$  para um número de graus de liberdade infinito a um nível de confiança de 95% ( $t_{c,0.05|N} = 1,96^3$ ). A hipótese nula assume que dois grupos, no caso deste trabalho por grupo entendase método de amostragem, não diferem significativamente. É rejeitada quando  $t_s$  é maior que o valor crítico  $t_c$ , sendo então admitida uma diferença significativa para os dois métodos de amostragem.

O grau de associação entre duas variáveis é observado a partir da análise da correlação. A correlação de Spearman baseia-se na ordenação de duas variáveis sem qualquer restrição quanto à distribuição de valores. O primeiro passo é a ordenação de uma variável e o segundo, a transformação dos valores absolutos em valores ordenados. As diferenças entre dois postos,  $d_i$ , são calculadas e o coeficiente de correlação,  $R$ , é determinado utilizando a Equação<sup>3</sup>:

$$R = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)} \quad (2)$$

Assume-se que uma associação é válida quando os coeficientes de correlação são menores que os valores críticos<sup>5</sup>.

A regressão de Kendall assume que uma série de pontos ( $X_1, Y_1$ ), ( $X_2, Y_2$ ) ... ( $X_n, Y_n$ ), é ajustada para uma recta com a forma:

$$Y = bX + a \quad (3)$$

análoga à regressão linear paramétrica, mas com ordenada na origem e coeficiente angular calculados a partir da transformação das variáveis ordenadas.

O primeiro passo consiste na ordenação dos valores da variável  $X$  por ordem crescente, para  $n$  observações. O declive,  $b_{ji}$ , é determinado para cada par de valores de  $X_i$  e  $X_j^4$ :

$$b_{ji} = \frac{Y_j - Y_i}{X_j - X_i} \quad (4)$$

em que  $i$  varia de 1 a  $n-1$  e  $j > i$ . O declive,  $b$ , é a mediana dos valores  $b_{ji}$ .

A ordenada de cada par de valores ( $X_i, Y_i$ ),  $a_i$ , é determinada utilizando a expressão<sup>4</sup>:

$$a_i = Y_i - bX_i \quad (5)$$

A ordenada na origem,  $a$ , é a mediana dos valores  $a_i$ .

Este método tem três vantagens distintas relativamente ao método paramétrico: não assume que todos os erros ocorrem na direcção de  $Y$ ; não assume que todos os erros são normalmente distribuídos e o resultado não é afectado pela presença de valores aberrantes.

## RESULTADOS E DISCUSSÃO

A relação entre as espécies iónicas em água do mar, isto é salinas, é conhecida e pode ser aplicada a aerossóis marinhos<sup>6</sup>. Considerando como elemento de referência o Na<sup>+</sup>, foi calculada a fracção não salina (nss) dos iões K<sup>+</sup>, Ca<sup>2+</sup> e SO<sub>4</sub><sup>2-</sup>:

$$\text{nss}W^i = W^i - \left[ \frac{W^i}{\text{Na}^+} \right]_{\text{salino}} \text{Na}^+ \quad (6)$$

sendo  $W$  a espécie e  $i$  a respectiva carga iónica.

Os testes não paramétricos foram efectuados para todas as espécies iónicas salinas e não salinas para comparação das metodologias de amostragem de alto e de baixo volume.

Estudos analíticos anteriores para a metodologia de elevado volume utilizando as duas metodologias de extracção (centrifugação e ultra-som) permitiram concluir que a metodologia analítica não conduz a diferentes resultados, permitindo afirmar que as diferenças observadas são devidas às diferentes metodologias de amostragem<sup>6</sup>.

Na Tabela 1 encontram-se os resultados da comparação das duas metodologias de amostragem, utilizando como exemplo comparativo de teste paramétrico a comparação emparelhada  $t$ -Student ( $t$ , calculado utilizando o programa Excel) e como métodos estatísticos não-paramétricos o teste de Wilcoxon ( $t_s$ ), o coeficiente de correlação de Spearman ( $R$ ) e a regressão de Kendall através da qual se obtiveram os factores de conversão para a metodologia de baixo volume em função da metodologia de alto volume. Os erros da regressão foram calculados considerando os desvios obtidos entre os valores experimentais da amostragem de baixo volume e os valores calculados a partir da conversão da metodologia de elevado volume em baixo volume.

Os valores estatísticos da comparação emparelhada  $t$ -Student e do teste de Wilcoxon (Tabela 1), são superiores ao valor crítico,  $t_c$ ,

**Tabela 1.** Resultados obtidos utilizando os métodos estatísticos para comparação das duas metodologias de amostragem relativamente a cada uma das espécies iónicas

Espécies Iónicas	<i>t-Student</i>	$t_s$ Wilcoxon	<i>R</i> Spearman	Regressão de Kendall	Erro Conversão
Na <sup>+</sup>	6,57	5,13	0,9443 <sup>(1)</sup>	$Y = 0,921 X + 22,097$	±0,25
NH <sub>4</sub> <sup>+</sup>	3,10	2,89	0,9159 <sup>(1)</sup>	$Y = 1,001 X + 0,627$	±0,27
K <sup>+</sup>	17,33	6,74	0,7064 <sup>(1)</sup>	$Y = 0,120 X + 0,205$	±0,34
Mg <sup>2+</sup>	3,22	2,93	0,9253 <sup>(1)</sup>	$Y = 0,815 X + 3,085$	±0,27
Ca <sup>2+</sup>	7,63	5,43	0,8518 <sup>(1)</sup>	$Y = 1,064 X - 1,159$	±0,29
Cl <sup>-</sup>	3,67	3,19	0,9176 <sup>(1)</sup>	$Y = 0,867 X + 31,046$	±0,27
NO <sub>3</sub> <sup>-</sup>	3,33	4,00	0,9866 <sup>(1)</sup>	$Y = 0,921 X + 0,036$	±0,23
SO <sub>4</sub> <sup>2-</sup>	10,22	6,62	0,9339 <sup>(1)</sup>	$Y = 0,840 X + 0,117$	±0,26
MSA	12,70	5,37	0,8392 <sup>(2)</sup>	$Y = 0,651 X - 0,102$	±0,30
nssSO <sub>4</sub> <sup>2-</sup>	10,16	6,57	0,9419 <sup>(3)</sup>	$Y = 0,741 X - 0,703$	±0,26
nssCa <sup>2+</sup>	9,88	6,05	0,8409 <sup>(1)</sup>	$Y = 0,891 X - 0,478$	±0,29
nssK <sup>+</sup>	15,95	6,74	0,5855 <sup>(1)</sup>	$Y = 0,113 X + 0,220$	±0,36

Valor crítico de  $t^2$ :  $t_{c,0,05|N|} = 1,96$ ; Valores críticos do coeficiente de correlação de Spearman (*R*), para *N* amostras<sup>5</sup>: <sup>(1)</sup> *N* = 60, *R* = 0,255; <sup>(2)</sup> *N* = 38, *R* = 0,321; <sup>(3)</sup> *N* = 57, *R* = 0,261; *X* = amostra de alto volume / nmol m<sup>-3</sup>; *Y* = amostra de baixo volume / nmol m<sup>-3</sup>

para todas as espécies iónicas em estudo. Verifica-se que tanto com o teste paramétrico como com o teste não-paramétrico se obtêm os mesmos resultados, ou seja, as duas metodologias de amostragem conduzem a resultados diferentes para todas as espécies.

K<sup>+</sup>, nssK<sup>+</sup>, Ca<sup>2+</sup> e nssCa<sup>2+</sup> são as espécies que apresentam maior dispersão de valores experimentais quando se comparam os resultados obtidos pelos dois métodos. Na Figura 1 encontra-se a representação gráfica dos valores experimentais da amostragem de baixo volume vs. amostragem de elevado volume e a recta da respectiva regressão de Kendall, para as espécies iónicas K<sup>+</sup> e Ca<sup>2+</sup>. O grau de associação das duas metodologias de amostragem para as espécies K<sup>+</sup>, nssK<sup>+</sup>, Ca<sup>2+</sup> e nssCa<sup>2+</sup>, utilizando os tradicionais métodos paramétricos não é válido. Utilizando métodos não-paramétricos já é possível efectuar regressão para determinação de factores de conversão.

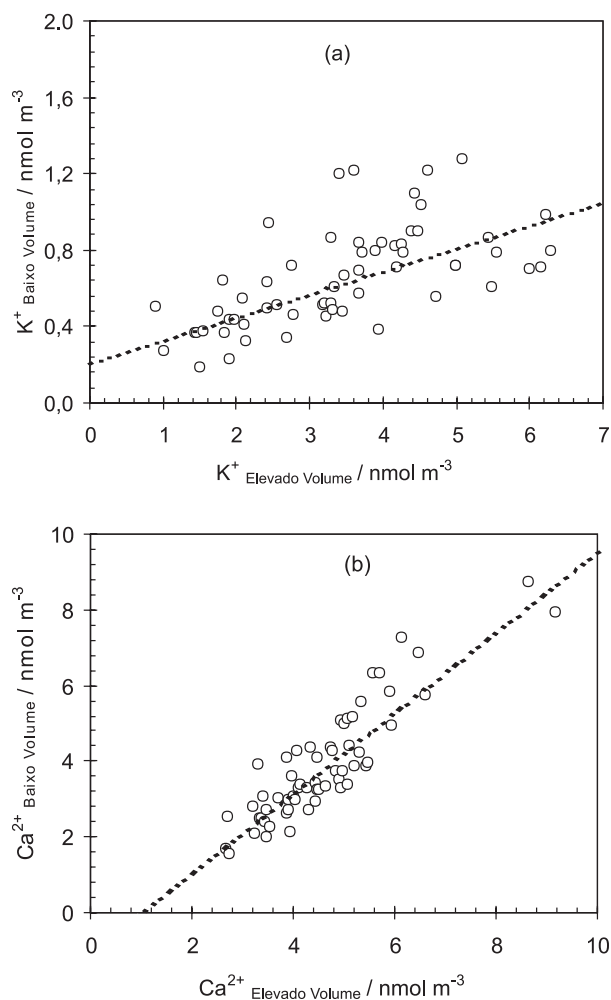
Todas as espécies iónicas em estudo apresentam um elevado grau de dependência e associação quando se comparam os dois métodos de amostragem, pois os valores calculados dos coeficientes de correlação de Spearman são mais elevados que os tabelados (Tabela 1).

Da regressão de Kendall (Tabela 1) verifica-se que todas as espécies iónicas têm factores de conversão diferentes, justificado pelos diferentes tamanhos das partículas e pelas suas diferentes propriedades físico-químicas.

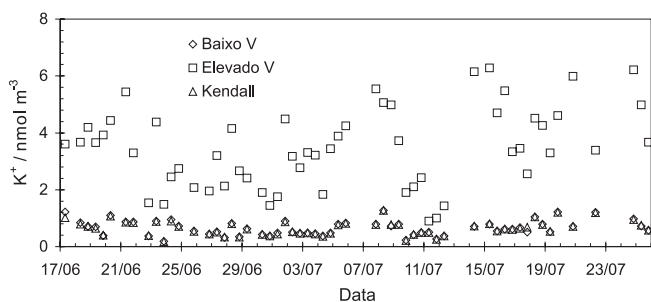
A título exemplificativo, na Figura 2 encontram-se as concentrações atmosféricas de K<sup>+</sup> obtidas para as duas metodologias de amostragem<sup>6</sup> e a conversão da metodologia de alto volume em baixo volume utilizando os factores de conversão da regressão de Kendall indicados na Tabela 1. O coeficiente de correlação de Pearson para a regressão linear é 0,6256 e para a logarítmica é 0,6991, motivo pelo qual os factores de conversão não podem ser calculados, pois a associação entre as duas metodologias não é considerada válida. Deste modo, através da representação gráfica, verifica-se experimentalmente a concordância entre as concentrações experimentais e as concentrações calculadas com teste não-paramétrico.

As concentrações atmosféricas<sup>6</sup> de aerossol NO<sub>3</sub><sup>-</sup> encontram-se representadas na Figura 3 e constituem um exemplo de uma espécie iónica em que os factores de conversão podem ser calculados utilizando estatística paramétrica ou não-paramétrica. Na zona de baixas concentrações, 0 a 10 nmol m<sup>-3</sup>, observa-se a concordância entre as concentrações experimentais de baixo volume e as calculadas para conversão da metodologia de elevado volume em baixo volume a partir dos factores de conversão de regressão linear, logarítmica e de Kendall. No entanto, na zona de concentrações mais elevadas, superiores a 20 nmol m<sup>-3</sup>, as concentrações calculadas utilizando os

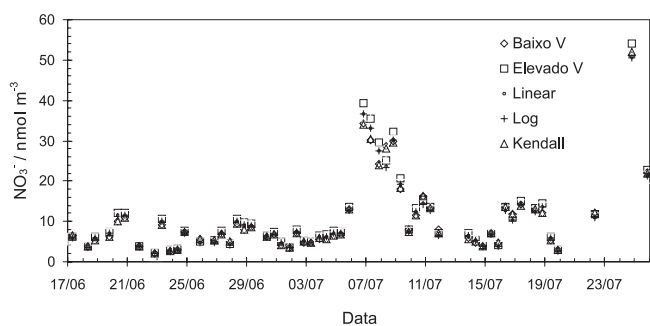
factores de conversão da regressão linear e logarítmica apresentam desvios mais elevados que as concentrações calculadas utilizando a regressão de Kendall. Os erros determinados utilizando os factores de conversão para a espécie NO<sub>3</sub><sup>-</sup> são de 1,15, 1,14 e 0,24 para a



**Figura 1.** Concentrações da amostragem de baixo volume vs. amostragem de elevado volume para as espécies iónicas que apresentaram maior dispersão de valores (a) K<sup>+</sup> e (b) Ca<sup>2+</sup>. Regressão de Kendall



**Figura 2.** Concentrações experimentais para a espécie iônica  $K^+$  com as duas metodologias de amostragem e as concentrações calculadas para a metodologia de baixo volume, utilizando os factores de conversão não-paramétrica da regressão de Kendall



**Figura 3.** Concentrações experimentais para a espécie iônica  $NO_3^-$  com as duas metodologias de amostragem e as concentrações calculadas para a metodologia de baixo volume, utilizando os factores de conversão paramétricos da regressão linear e logarítmica e os factores de conversão não-paramétrica da regressão de Kendall

regressão linear, logarítmica e regressão de Kendall, respectivamente. Verifica-se, deste modo, que o teste não-paramétrico é o que apresenta um erro menor, embora os erros obtidos pelos outros testes não invalidem os mesmos.

Os erros calculados para a regressão de Kendall (Tabela 1) são mais elevados quando os graus de associação das variáveis, determinados utilizando a correlação de Spearman, são menores e inferiores a 5%.

Tradicionalmente, em Química da Atmosfera, são utilizados métodos paramétricos para comparação de métodos. Neste trabalho verifica-se que a conversão de resultados de diferentes metodologias de amostragem com ferramenta não-paramétrica é mais adequada, pois mesmo no caso em que se obtém maior dispersão de resultados torna possível a comparação.

Estudos futuros poderão verificar se estes factores se alteram ou se mantêm, pois diferentes condições meteorológicas e diferentes emissões antropogénicas, biogénicas e minerais deverão conduzir a diferentes propriedades químico-físicas dos aerossóis.

## CONCLUSÃO

O teste de Wilcoxon permite concluir que os resultados obtidos para cada espécie iônica são diferentes para cada uma das metodologias de amostragem. No estudo apresentado, o teste *t*-Student pode ser aplicado com sucesso para todas as espécies iônicas.

Um elevado grau de associação para as duas metodologias de amostragem foi determinado utilizando a correlação de Spearman, o que permitiu o cálculo dos factores de conversão através da regressão de Kendall.

Resultados comuns às duas metodologias para todas as espécies em estudo, seria a situação ideal. No entanto, as concentrações atmosféricas dos aerossóis para cada espécie iônica resultam diferentes, como consequência de o ar amostrado se encontrar em movimento e de a velocidade de sucção do ar ser diferente. Os factores de conversão para cada espécie iônica foram determinados para metodologia de baixo volume em função da metodologia de alto volume.

## AGRADECIMENTOS

M. J. Nunes agradece à Fundação da Ciência e Tecnologia, Projecto POCTI/CTA/36269/99 e ao Joint Research Centre EC, Ispra, pelo suporte financeiro logístico à realização experimental deste trabalho.

## REFERÊNCIAS

1. <http://www.stawds.strath.ac.uk/~eillen/classes/53904/mod2904>, acessada em Março 2003.
2. <http://www.sportsci.org/resource/stats/nonparms.htm>, acessada em Março 2003.
3. Miller, J. C.; Miller, J. N.; *Non-parametric and robust methods in Statistics for Analytical Chemistry*, 3<sup>rd</sup> ed., Ellis Horwood PTR Prentice Hall: London, 1993.
4. Sokal, R. R.; Rohlf, E. J.; *The normal probability distribution, Estimation and hypothesis testing, Introduction to the analysis of variance, Assumptions of analysis of variance, Linear regression, Correlation in Biometry*, 3<sup>rd</sup> ed., W. H. Freeman and Company: New York, 1995.
5. Neave, H. R.; *Statistics Tables for mathematicians, engineers, economists and the behavioral and management sciences*, George Allen & Unwin: London, 1978.
6. Nunes, M. J.; *Tese de Doutoramento*, Universidade de Lisboa, Portugal, 2002.